Spring 2014

The Production Effect and Item-Order Encoding

Bethany Kregiel
John Carroll University, bkregiel15@jcu.edu

The Production Effect and Item-Order Encoding

Bethany Kregiel

John Carroll University

Abstract

When reading a mixed list of words, participants show better memory for uncommon words compared to common words (McDaniel & Bugg, 2008). The research suggests differential memory effects in item-order encoding between mixed and pure lists. Uncommon words lead to item-specific encoding whereas common words lead to order encoding. Similarly, the production effect shows that, when reading a mixed list (some words aloud, others silently), participants show better memory for the words read aloud, but the effect does not obtain for pure lists. The purpose of this study is to examine if the production effect is due to differences in item-order encoding. Sixty-five John Carroll University undergraduates read six lists of sixteen words one at a time. Some participants read all words aloud (pure aloud), some read all words silently (pure silent); some read half of the words aloud and half silently depending on font color (mixed). At the end of each list, all participants completed a one-minute free recall task. After the final free recall task for the last list, all participants completed an order reconstruction task. Recall accuracy, input-output correspondence, and order reconstruction were examined using ANOVAs and t-tests. Basic production effect findings were replicated; aloud words were better remembered than silent words only for the mixed list group (Jones & Pyc, 2013). Further, because aloud words can be considered "uncommon", we saw a decrease in order measures for the aloud items on a mixed list compared to the pure list. Similarly, order measures increased for silent words, which can be considered "common". Thus, the production effect can be considered another example of how item-order encoding varies in mixed/pure list learning.

The Production Effect and Item-Order Encoding

The production effect is the memorial benefit for items read aloud compared to items read silently in a mixed list (Bodner & Taikh, 2012; Bodner, Taikh & Fawcett, 2014; Jones & Pyc, 2014; Jonker, Levene, & MacLeod, 2013; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Ozubko & MacLeod, 2010). In other words, when given a list of items, if one reads half of the items aloud and half of the items silently, one will have better recall for the items read aloud. Importantly, this effect does not occur in a pure list, or rather, when one reads all items aloud or all items silently.

Recent research on the production effect has revealed that, while there appears to be a memorial benefit for aloud items in a mixed list compared to a pure list, a memorial cost occurs for silent items in a mixed list compared to a pure list (Bodner et al., 2014; Jones & Pyc., 2014). Thus, one would have better memory for silent items if all of the items had been read silently, yet one would have better memory for aloud items if half of the items had been read aloud and half had been read silently. Jones and Pyc (2014) found that the production effect results in a cost to silent items and a benefit to aloud items using a free recall task, which required the participants to list all items that he or she could remember from the originally studied list. Bodner et al. (2014) replicated similar findings using a recognition task, which prompted the participant to decide whether or not an item that they were given had been viewed previously on a list that they had studied. Overall, the production effect appears to result in a memorial difference based on pure lists and mixed lists, which begs the question: why is there a memorial difference based on list type? Recent research has evaluated two different explanations for the production effect: the strength theory and the distinctiveness theory.

According to the strength theory, the production effect occurs because the encoding strength differs for items read aloud and items read silently (Bodner & Taikh, 2012). Specifically, increased memory for items read aloud may result from greater strength of encoding for aloud words compared to silent words. This theory suggests that items read aloud lead to stronger encoding, making those items easier to retrieve from memory than the silent items. Tests of the strength theory require the participant to read both a pure list and a mixed list to distinguish whether or not the fact that word is spoken aloud leads to better memory for those items regardless of list type. However, the results regarding the strength theory have been inconclusive (Bodner & Taikh, 2012).

Another explanation, the distinctiveness theory, states that items read aloud result in better memory due to the fact that they are distinct from silent items (Ozubko & MacLeod, 2010). In other words, items read aloud stand out because one can remember verbalizing those words when other words are not vocalized. For example, when taking a recognition test, a participant can say of a word read aloud, “I remember speaking that word; therefore, it was on the list.” However, there is nothing distinct about a word read silently. The participant cannot remember voicing the word, and therefore, he or she has a harder time recognizing it on a recognition test. This theory would explain why there is benefit for items read aloud and a cost for items read silently in a mixed list.

Interestingly, by examining effects similar to the production effect (i.e., effects in which memorial differences occur between mixed lists and pure lists), a third explanation emerges. Consider the generation effect, which is a memorial benefit for items generated by a participant compared to items given to a participant by the experimenter (McDaniel & Bugg, 2008). For example, a participant will have better memory for a word that they were asked to generate from a word stem compared to a word that the experimenter gave to them to study in a list. Similarly, the word frequency effect is a memorial benefit for high-frequency words (words that one encounters frequently) compared to low-frequency words (words that one encounters rarely) (DeLosh & McDaniel, 1996). Another example refers to better memory for humorous items compared to non-humorous items (Schmidt, 1994). All of these effects mirror the production effect in that whether or not they occur depends upon whether they are part of a mixed or pure list. Thus, examination of the causes of these effects can provide insight into the production effect.

It has been proposed that the memorial differences for the aforementioned effects occur due to differences in memory encoding based on the item type (McDaniel & Bugg, 2008). Previous literature suggests that phenomena such as the generation effect, the frequency effect, or the humor effect occur due to encoding differences for common items and uncommon items. Specifically, common items lead to encoding based on the order in which it appeared in the list, whereas uncommon items lead to encoding based on specific information about that item. Consider, once again, the frequency effect. According to the item-order encoding theory, high-frequency items would be considered common, and thus, would lead to encoding based on order. Therefore, when an individual would attempt to retrieve a common item from memory, he or she would focus on the serial order in which that item had appeared in the list (i.e., “I remember the word “author” because it came after the word “partner”). Contrastingly, the item-order encoding theory would consider low-frequency items to be uncommon. Thus, these uncommon items would lead to encoding based on item-specific information, and an individual would retrieve that item based on the fact that it is a low-frequency item (i.e., “I remember the word “turnip” because it is a word that I rarely encounter”). Importantly, these encoding differences tend to vary based on list design. For instance, uncommon items tend to get better order encoding when in a mixed list, whereas common items tend to get poorer order encoding when in a mixed list (McDaniel & Bugg, 2008).

This same theory may be a possible explanation for the production effect. Words read silently and words read aloud differ in commonness. Specifically, words read silently are considered common, and words read aloud are considered uncommon. It is typical, for example, that one would read the newspaper or a book silently to oneself as opposed to reading it aloud. However, when an individual reads a word aloud, this item would be processed with an auditory cue, making it unlike the silent item and thus, uncommon. Therefore, if the production effect maps onto the item-order encoding theory, words read silently should lead to order encoding (i.e., "I remember the word "author" because it came after the word "partner") and words read aloud should lead to item-specific encoding (i.e., "I remember the word "turnip" because I remember reading it aloud).

The item-order encoding theory has been evaluated in regard to the production effect in one recent study (Jonker, Levene & MacLeod, 2014). To test the item-order encoding theory, Jonker et al. (2014) randomly assigned participants into one of three conditions: one in which they read all words aloud, one in which they read all words silently, and one in which they read half of the words aloud and half of the words silently. In each of these conditions, participants were instructed to study the eight items presented, and upon completion of the study portion, the participants completed a short distractor activity. After the distractor activity, the participants were prompted with an order reconstruction task half of the time and a free recall task the other half of the time. For the reconstruction task, the participants were presented with all items that had previously been shown in the list, and they were asked to put these items back into the order in which they were originally studied. For the free recall task, participants were asked to list all of the items they recalled studying in the order in which they came to mind. The purpose of both of these tasks was to evaluate the degree to which the participants encoded items based on serial order information. Thus, the researchers predicted results consistent with the item-order encoding theory: silent items would lead to the best order encoding and aloud items would lead to the worst order encoding for pure lists; order encoding for mixed lists would fall in the middle. Overall, the results from both tasks showed that silent items did lead to better order encoding compared to aloud items, with the suggestion that aloud items led to item-specific encoding

Although these results are consistent with predictions, it is important to note methodological issues that may have influenced results. First, the aforementioned study presented participants with lists of eight items. However, other production effect research has evaluated the effect using more than eight items, ranging anywhere from sixteen to eighty (Bodner et al., 2012; Bodner et al., 2014; Jones & Pyc., 2014; MacLeod et al, 2010; Ozubko et al., 2010). Additionally, the previous item-order and production effect research presented participants with a reconstruction task after half of the lists (Jonker et al., 2014). While this type of task is successful in evaluating order encoding of certain items, their particular design may have biased the results in that the participants were expecting the reconstruction task, which may have prompted participants to pay greater attention to serial order when first memorizing the lists.

In contrast, the current study attempts to amend these two methodological issues. First, participants in this experiment will study a list of sixteen items instead of eight items, which should align the results more accurately with previous production effect research. Additionally, the current research presents the participants with a spontaneous reconstruction task as the final part of the experiment instead of one after half of the lists. This unexpected measure should provide a more natural evaluation of order encoding, accounting for any bias that may have occurred in the previous study. As such, the purpose of the current study is to investigate the item-order encoding theory as an explanation for the production effect using stronger methodology than the aforementioned study.

Method

Participants

Sixty-five John Carroll University students participated in exchange for course credit.

Materials and Procedures

The participants were randomly assigned to one of three conditions: mixed list, pure aloud list, or pure silent list. Each participant, regardless of condition, was given six lists of sixteen words. Each word was shown individually for two seconds at a time. The words were randomly presented in either an orange font or a green font to counterbalance the amount of words a participant would read aloud or silently. For the mixed lists, the color of the font indicated which words should be read aloud (i.e., “If the word is orange, read it aloud”) and which words should be read silently (i.e., “If the word is green, read it silently”). For the pure lists, participants either read all words aloud (pure aloud) or all words silently (pure silent), regardless of the font color. After each list, participants completed a thirty second distractor activity. Following the distractor task, participants were prompted to complete a free recall task by typing in all words that they could remember from the previous list. Participants were given one minute to complete the free recall task. For the final list of sixteen words, participants were asked to complete an unexpected order reconstruction task following the free recall task. They were shown all of the items from the last list in a randomized order and were instructed to place them in the order in which they were originally studied.

Results

Because this particular study was intended to create a more rigorous evaluation of the production effect and item-order encoding, the results were evaluated in four different ways. First, to ensure replication of basic production effect findings, recall accuracy was evaluated by comparing the mean percentages of words recalled in both mixed and pure lists. Second, input-output correspondence was evaluated in two ways, one based on relative order and the other based on the average distance between the words recalled. Finally, order reconstruction was evaluated based on the accuracy of the order reconstruction task given to the participants in the last portion of the experiment. These forms of evaluation will be discussed in greater detail in the sections with the respective titles.

Free Recall Accuracy

Although not the primary focus of the current study, recall accuracy was examined to assess production effect replication. Recall accuracy was evaluated using a 2 (production: aloud, silent) x 2 (list type: mixed, pure) mixed factor Analysis of Variance (ANOVA). As is common within production effect literature (Ozubko et al., 2010; Jones & Pyc, 2014), production was treated as a within-subjects manipulation, and list type was treated as a between-subjects manipulation. There was a main effect of production, $F(1, 41) = 27.002$, $p < .001$, $\eta^2 = .397$, showing overall greater recall for words read aloud ($M = .354$, $SD = .132$) compared to words read silently ($M = .237$, $SD = .117$). There was no effect of list type, $F(1, 41) = 2.220$, $p = .144$, $\eta^2 = .051$. More importantly, there was an interaction between production and list type, $F(1, 41) = 20.573$, $p < .001$, $\eta^2 = .334$. The interaction indicated that there were no significant differences between pure lists, $t(21) = .448$, $p = .630$. However, as shown in Figure 1, recall for aloud items was greater than recall for silent items in a mixed list, $t(20) = 6.590$, replicating basic production effect findings. When comparing mixed and pure lists focusing on silent items alone, there was a significant difference for silent items, $t(41) = -4.988$, $p < .001$, showing that recall accuracy was greater for silent items in a pure list ($M = .307$, $SD = .094$) compared to a mixed list ($M = .165$, $SD = .092$). Alternatively, when comparing mixed and pure lists focusing on aloud items alone, the difference did not reach significance, $t(41) = 1.675$, $p = .101$, but the means for recall showed a trend toward greater accuracy for aloud items in a mixed list ($M = .388$, $SD = .141$), compared to a pure list ($M = .322$, $SD = .117$). Overall, these results replicate prior work showing costs to silent items and benefits to aloud items from a pure list to a mixed list (Bodner et al., 2014; Jones & Pyc, 2014).

Order Encoding

Relative input-output correspondence. Input-output correspondence was evaluated in one of two ways. The first way, referred to in this paper as relative input-output correspondence, has been used in previous item-order encoding literature (DeLosh & McDaniel, 1996) to evaluate input-output correspondence. The evaluation system was based off of a scoring system originally created by Asch and Ebenholtz (1962), and it grants a participant a point for input-output correspondence based on the proportion of items recalled in a forward moving order.

For example, consider a participant who recalled the following items (shown with the respective serial position following the item): nephew (1), distance (2), garden (3), turnip (11), shoulder (5), sailor (7), partner (6), author (14). Overall, there were five items recalled in a forward moving direction: (nephew to distance (1); distance to garden (2), garden to turnip (3), shoulder to sailor (4), and partner to author (5). The participant would not receive a relative input-output correspondence score for the items "turnip" and "partner," because these items were recalled in a backward moving order, from serial positions 11 to 5 and 7 to 6, respectively. The final relative input-output correspondence score was obtained by taking the amount of items recalled in a forward moving direct (in this example, five), divided by the total number of pairs recalled (in this example, seven). Therefore, the participant in the example would receive a final score of approximately .742. In terms of relative input-output correspondence scores, the higher the score, the better the input-output accuracy. Refer to Table 1 for a visual representation of this example.

Based on this scoring system, there was an effect of list type, $F(2, 62) = 3.531$, $p = .035$, $\eta^2 = .102$, with a significant difference between pure aloud lists and pure silent lists, $p = .010$. As shown in Figure 2, this difference showed greater input-output correspondence for pure silent lists ($M = .587$, $SD = .024$) compared with pure aloud lists ($M = .495$, $SD = .024$). While there was not a significant difference between pure aloud lists and mixed lists, $t(41) = 1.642$, $p = .142$, the results suggest that input-output correspondence is greater in mixed lists ($M = .547$, $SD = .025$) compared to pure aloud lists ($M = .495$, $SD = .024$). Thus, the relative input-output correspondence was greatest in pure silent lists, then mixed lists, then pure aloud lists. These results are consistent with the predictions.

Distance. The second evaluation of input-output correspondence uses a scoring system based from Jonker et al. (2013). Instead of scoring participants based on items moving in only a forward direction, this scoring system determined an average distance between items, regardless of the direction in which they were recalled.

For example, consider that same participant who had recalled the following items (shown with the respective serial position following the item): nephew (1), distance (2), garden (3), turnip (11), shoulder (5), sailor (7), partner (6), author (14). The first part of this scoring system requires calculating the distance between each item. This participant would receive the following points for distance: 1 for the distance between items one and two, 1 for the distance between items two and three, 8 for the distance between items three and eleven, 6 for the distance between items eleven and five, 2 for the distance between items five and seven, 1 for the distance between items seven and six, and 8 for the distance between items six and fourteen. These points would then be added together for a total distance of 27. The total point score would then be divided by the number of pairs recalled—in this case, 7—for an average distance of approximately 3.9 between recalled items. In terms of distance input-output correspondence, the lower the score, the better the order encoding. Once again, refer to Table 1 for a visual representation of this example.

Based on this scoring system for input-output correspondence, an effect of list type was found, $F(2, 62)$, $p = .031$, $\eta^2 = .106$. This effect showed a significant difference in distance scores for pure aloud lists and pure silent lists, $t(42) = 2.162$, $p = .036$, with lower distance scores for pure silent lists ($M = 4.527$, $SD = .180$) compared with pure aloud lists ($M = 5.063$, $SD = .180$), as shown in Figure 3. This effect also showed a significant difference between pure aloud lists and mixed lists, $t(41) = 2.665$, $p = .011$, with, once again, lower distance scores for mixed lists ($M = 4.411$, $SD = .185$) compared to pure aloud lists ($M = 5.063$, $SD = .180$). Thus, much like relative input-output correspondence, the input-output correspondence based on distance was greatest for pure silent lists, then mixed lists, then pure aloud lists.

Order reconstruction task. The order reconstruction task completed by participants at the end of the experiment is the final and strictest dependent measure. It is important to note that participants only received a score for order reconstruction if, during the reconstruction task, they placed an item in the *exact serial position* as the position it had been studied in the original list. For example, imagine that the item “turnip” appeared as the eleventh item in the original list of sixteen items. The participant would only receive a point for reconstruction if he or she placed the item “turnip” in the eleventh position on the reconstruction task. He or she would not receive any points if the item “turnip” appeared in any other position than the eleventh position for the reconstruction task.

To evaluate reconstruction, a 2 (production: aloud, silent) x 2 (list type: mixed, pure) mixed factor Analysis of Variance (ANOVA) was used. There was no main effect of production, $F(1, 37) = 1.092$, $p = .303$, $\eta^2 = .029$. Additionally, there was no interaction between production and list type $F(1, 37) = .246$, $p = .623$, $\eta^2 = .007$. However, there was an effect of list type, $F(1, 37) = 4.952$, $p = .032$, $\eta^2 = .118$.

When evaluating reconstruction for mixed lists only, t-tests showed no significant difference between aloud items and silent items, $t(19) = .497$, $p = .625$, as shown in Figure 4. Similarly, when evaluating reconstruction for pure lists only, t-tests showed no significant difference between aloud items and silent items, $t(18) = .911$, $p = .375$.

When evaluating reconstruction of aloud items and silent items in a mixed list compared to aloud items and silent items in a pure list, no significant difference was found, $t(39) = -1.175$, $p = .247$. Finally, when evaluating reconstruction of silent items in a mixed list compared to a pure list, no significant difference was found, $t(37) = -1.036$, $p = .307$.

Discussion

The current research investigated the production effect in regard to item-order encoding and proposed that the memorial difference observed in the production effect can be attributed to differences in encoding. Specifically, the current study suggested that silent items, which are considered common, lead to order encoding, whereas aloud items, which are considered uncommon, lead to item-specific encoding. The results of the current study replicated previous production effect findings while also showing that order encoding differs for aloud and silent items in mixed and pure lists.

First, it is important to note that the current study replicated basic production effect findings. As expected, participants showed greater recall for aloud items in a mixed list compared to silent items; thus, the production effect occurred. Further, the results also replicated previous findings which suggest that, while the production effect seems to produce a memorial benefit for items read aloud, it also produces a memorial cost for items read silently (Bodner et al., 2014; Jones & Pyc, 2014).

Of greater importance, however, both evaluations of input-output correspondence yielded results consistent with the predictions regarding item-order encoding. Specifically, silent items yielded the best scores for input-output correspondence, suggesting that silent items lead to order encoding, as had been predicted. Additionally, aloud items yielded the worst input-output correspondence scores, suggesting that aloud items do not lead predominantly to order encoding. Currently, there is no measure in this study or in other studies that evaluates item-specific encoding. However, because the results suggest that participants do not tend to retrieve aloud items based on order information, one can assume that aloud items lead to item-specific encoding. It is also important to note that input-output correspondence scores were worse in mixed lists compared to pure silent lists but better in mixed lists compared to pure aloud lists. Thus, it seems as if encoding changes from mixed lists to pure lists. Overall, the results from the current research align with results from previous research evaluating item-order encoding phenomena, such as the generation effect, the word frequency effect, and the humor effect

(McDaniel & Bugg., 2008; DeLosh et al., 1996; Schmidt, 1994) in that silent (i.e., common) items lead to order encoding and aloud (i.e., uncommon) items lead to item-specific encoding.

Some of the results from this experiment regarding input-output correspondence are consistent with previous production effect and item-order encoding research. Importantly, the current study replicated results from Jonker et al. (2013) in regard to distance scoring for input-output correspondence. In the aforementioned study, the results showed the best input-output correspondence for pure silent lists, then for mixed lists, and lastly, the worst input-output correspondence for aloud lists. Additionally, the current study replicated results from Jonker et al. (2013) in regard to relative input-output correspondence. Both studies showed that relative input-output correspondence was best for pure silent lists, then mixed lists, then aloud lists. Thus, the results from both studies add support to the item-order encoding theory for the production effect.

Although the current study did not find results consistent with predictions in regard to the reconstruction task, some methodological issues may explain why the current results differ from results found in previous the only other investigation of item-order encoding and the production effect (Jonker et al., 2014). As discussed previously, the current study presented participants with lists consisting of sixteen items, compared to eight items. Because the current study doubled the number of items from previous research, participants may have struggled to reconstruct the lists as accurately as they may have been able to if they had been prompted with only eight items instead. Another methodological difference refers to the way in which participants received the order reconstruction task. For instance, in Jonker et al. (2014), participants completed a reconstruction task after half of the lists. In the current study, participants only completed a reconstruction task after the final list, similar to designs from other studies investigating item-order encoding with other memory effects (Merritt et al., 2006). Thus, Jonker et al. (2014) may have found a significant effect because participants knew they would complete a reconstruction task and, in turn, they paid greater attention to order than they would have if the reconstruction task had been unexpected. Finally, it is important to note that the scoring for the reconstruction task was very strict. Participants could only receive a point for reconstruction if they placed the item in the *exact* serial order that it had appeared in on the list. This is incredibly difficult, especially for a list of sixteen items. Thus, perhaps future research could evaluate reconstruction using a scoring system that is less strict.

Future production effect and item-order encoding research is currently underway and will evaluate the same measures using lists of only eight items instead of sixteen. This methodological change may yield results consistent with previous research in regard to the order reconstruction task. Data collection and scoring is currently in progress.

In conclusion, the item-order encoding theory seems to be an adequate explanation for the production effect. The memorial differences seem to occur due to differences in encoding. Specifically, silent items lead to order encoding, and aloud items lead to item-specific encoding, thus making aloud items easier to remember in a mixed list.

References

Asch, S. E., & Ebenholtz, S. M. (1962). The process of free recall: Evidence for non-associative factors in acquisition and retention. *The Journal Of Psychology: Interdisciplinary And Applied, 54*(1), 3-31.

Bodner, G. E., & Taikh, A. (2012). Reassessing the basis of the production effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1711-1719.

Bodner, G. E., Taikh, A., & Fawcett, J.M. (2014). The costs and benefits of production in recognition. *Psychonomic Bulletin & Review.*

DeLosh, E.L., & McDaniel, M. A. (1996). The role of order information in free recall: Application to the word-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1136-1146.

Gregg, V.H., & Montgomery, D.C. (1980). Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior, 19,* 240-245.

Jones, A. C., & Pyc, M. A. (2014). The production effect: Costs and benefits in free recall. *Journal Of Experimental Psychology: Learning, Memory, And Cognition, 40*(1), 300-305.

Jonker, T. R., Levene, M., & MacLeod, C. M. (2014). Testing the item-order account of design effects using the production effect. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*, *40*(2), 441-448.

McDaniel, M.A., & Bugg, J.M. (2008). Instability in memory phenomena: a common puzzle and a unifying explanation. *Psychonomic Bulletin & Review, 15,* 237-255.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 671-685.

Merritt, P.S., DeLosh, E.L., & McDaniel, M.A. (2006). Effects of word frequency on individual-item and serial order retention: Tests of the order-encoding view. *Memory & Cognition, 34,* 1615-1627.

Ozubko, J.D., & MacLeod, C.M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1543-1547.

Schmidt, S.R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 953-967.

Table 1

Original List	**Relative Input-Output Correspondence**	**Distance Input-Output Correspondence**
1. nephew 2. distance 3. garden 4. fashion 5. shoulder 6. partner 7. sailor 8. village 9. holiday 10. leather 11. turnip 12. ticket 13. wagon 14. author 15. steam 16. journey	1. nephew (1) 2. distance (2) 3. garden (3) 4. turnip (11) 5. shoulder (5) 6. sailor (7) 7. partner (6) 8. author (14) Score: 5 Total Pairs: 7 Total Score: .742	1. nephew (1) 2. distance (2) 3. garden (3) 4. turnip (11) 5. shoulder (5) 6. sailor (7) 7. partner (6) 8. author (14) Total Distance: 27 Total Pairs: 7 Average Distance: 3.9

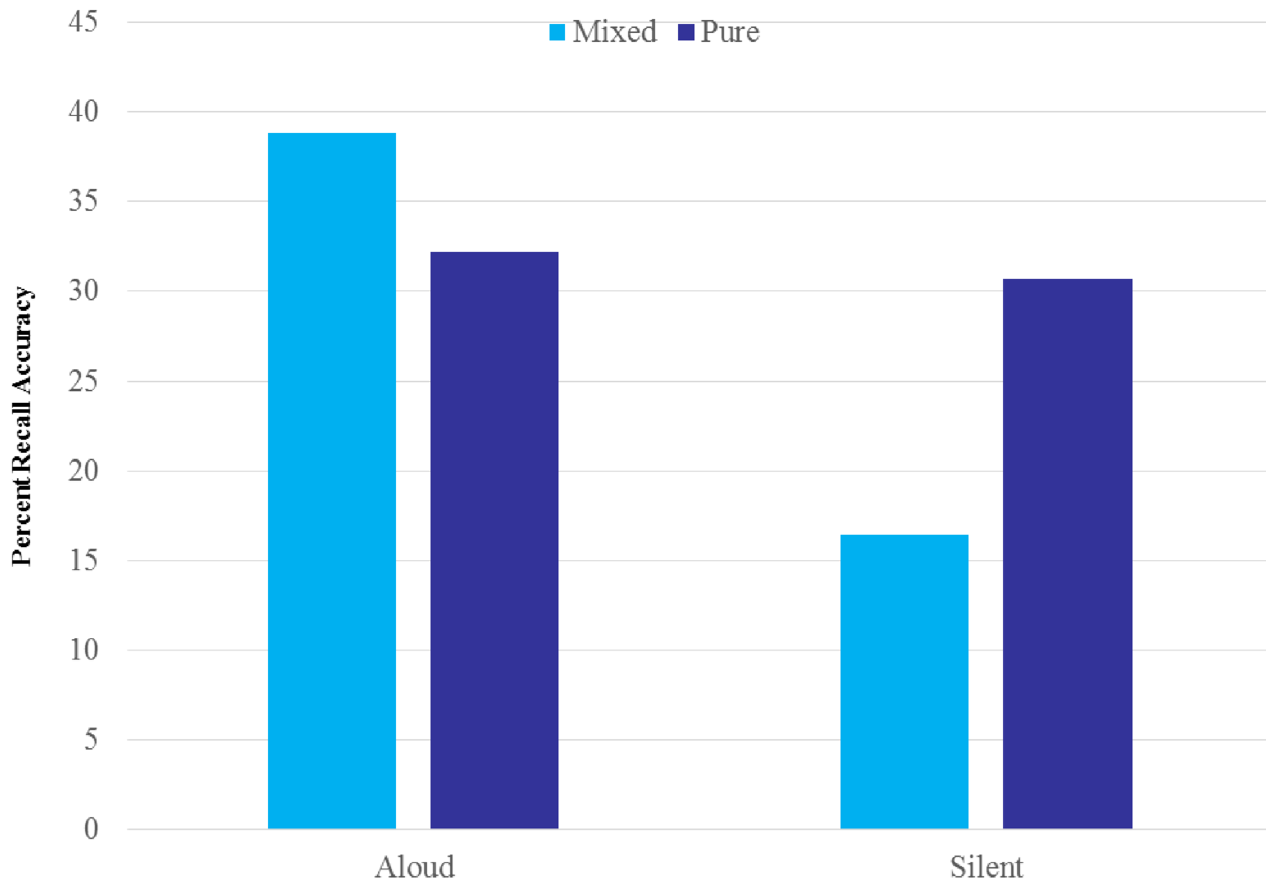

Figure 1. Free Recall Accuracy (Percent of items recalled in mixed and pure lists)

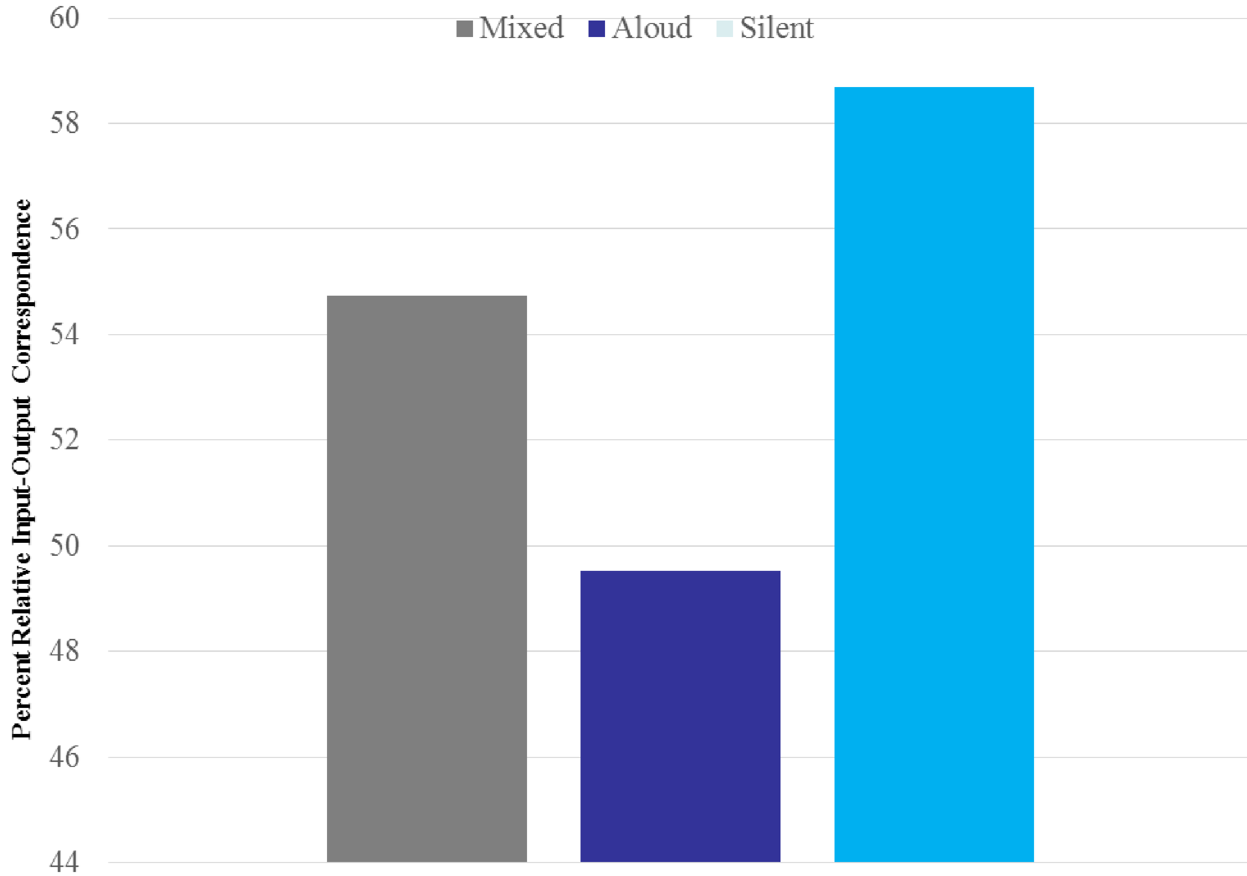

Figure 2. Relative Input-Output Correspondence (Percent of relative input-output scores for mixed lists, pure aloud lists, and silent lists)

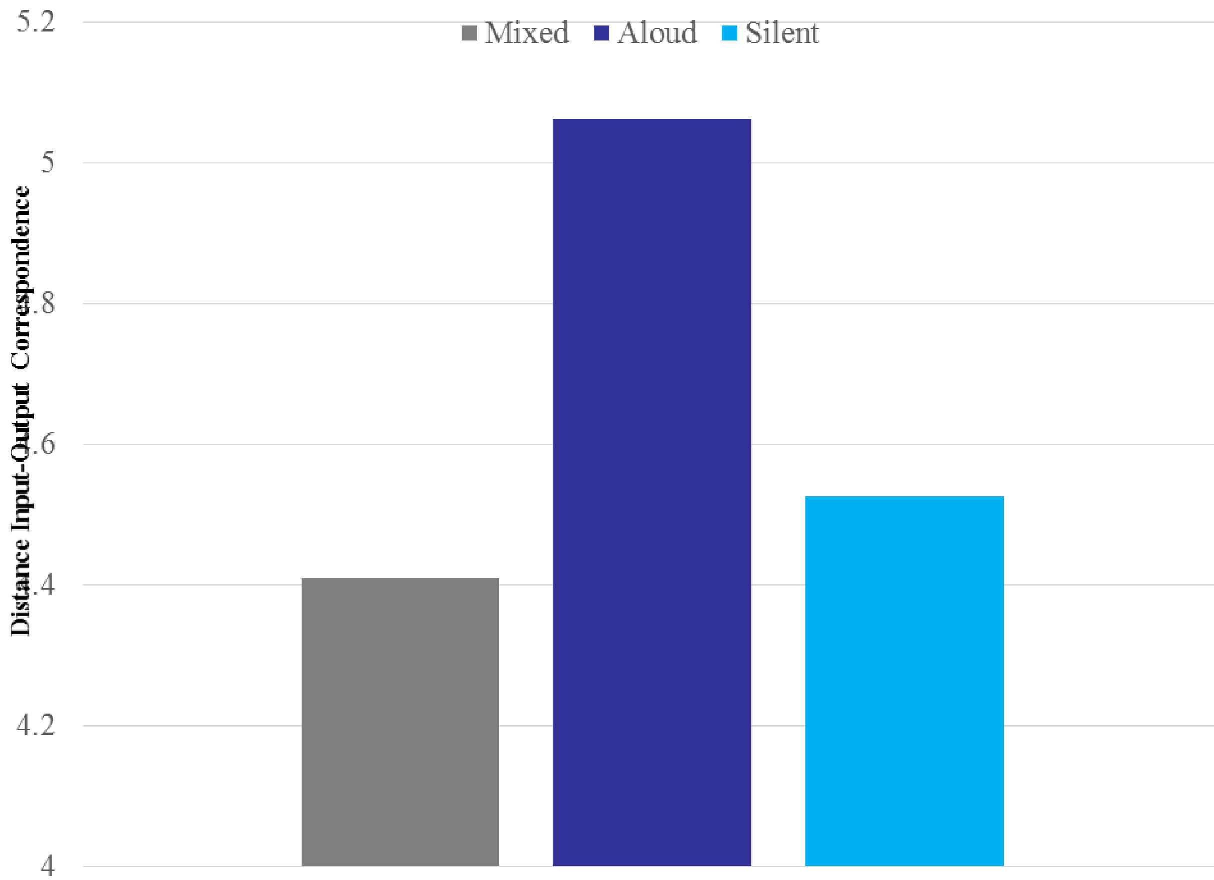

Figure 3. Distance Input-Output Correspondence (Average distance between items recalled in mixed lists, pure aloud lists, and pure silent lists)

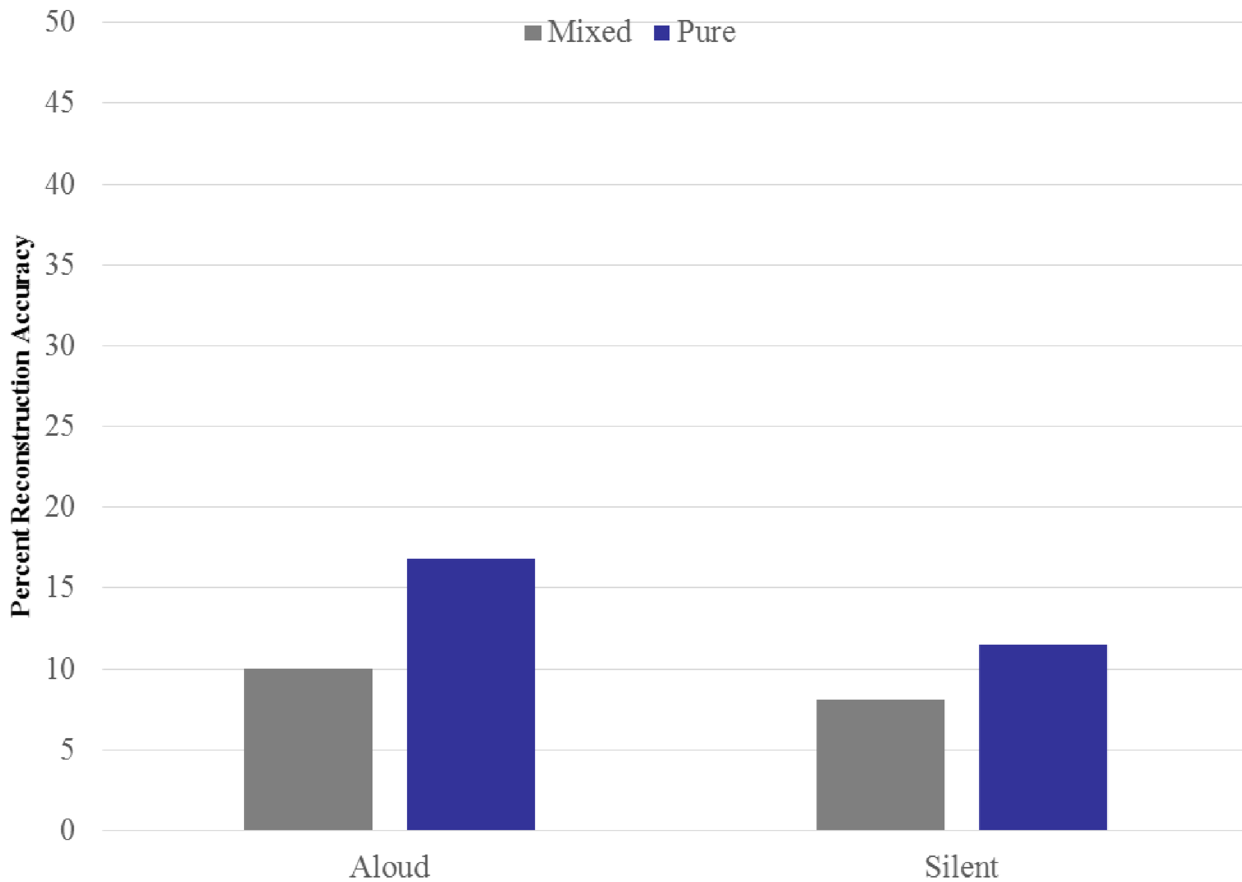

Figure 4. Percent Reconstruction Accuracy (Percent of items accurately reconstructed in mixed and pure lists)